

## **RAW SEQUENCE LISTING**

**The Biotechnology Systems Branch of the Scientific and Technical  
Information Center (STIC) no errors detected.**

Application Serial Number: 10/580,872  
Source: IFWP  
Date Processed by STIC: 6/22/06

# ***ENTERED***



IFWP

## RAW SEQUENCE LISTING

DATE: 06/22/2006

PATENT APPLICATION: US/10/580,872

TIME: 12:53:02

Input Set : A:\SEQUENCE LISTING.txt

Output Set: N:\CRF4\06222006\J580872.raw

```

3 <110> APPLICANT: Gaisser, Sabine
4 Haydock, Stephen F.
5 Leadlay, Peter, F.
6 McArthur, Hamish A.I.
8 <120> TITLE OF INVENTION: Polyketides and their synthesis
10 <130> FILE REFERENCE: 0380-P04095US00
C--> 12 <140> CURRENT APPLICATION NUMBER: US/10/580,872
C--> 13 <141> CURRENT FILING DATE: 2006-05-25
15 <150> PRIOR APPLICATION NUMBER: PCT/GB2004/005001
16 <151> PRIOR FILING DATE: 2004-11-29
18 <150> PRIOR APPLICATION NUMBER: GB0327721.7
19 <151> PRIOR FILING DATE: 2003-11-28
21 <160> NUMBER OF SEQ ID NOS: 57
23 <170> SOFTWARE: PatentIn version 3.2
25 <210> SEQ ID NO: 1
26 <211> LENGTH: 305
27 <212> TYPE: PRT
28 <213> ORGANISM: Streptomyces fradiae
30 <400> SEQUENCE: 1
32 Met Asn Asp Arg Pro Arg Arg Ala Met Lys Gly Ile Ile Leu Ala Gly
33 1 5 10 15
36 Gly Ser Gly Thr Arg Leu Arg Pro Leu Thr Gly Thr Leu Ser Lys Gln
37 20 25 30
40 Leu Leu Pro Val Tyr Asp Lys Pro Met Ile Tyr Tyr Pro Leu Ser Val
41 35 40 45
44 Leu Met Leu Ala Gly Ile Arg Glu Ile Gln Ile Ile Ser Ser Lys Asp
45 50 55 60
48 His Leu Asp Leu Phe Arg Ser Leu Leu Gly Glu Gly Asp Arg Leu Gly
49 65 70 75 80
52 Leu Ser Ile Ser Tyr Ala Glu Gln Arg Glu Pro Arg Gly Ile Ala Glu
53 85 90 95
56 Ala Phe Leu Ile Gly Ala Arg His Ile Gly Gly Asp Asp Ala Ala Leu
57 100 105 110
60 Ile Leu Gly Asp Asn Val Phe His Gly Pro Gly Phe Ser Ser Val Leu
61 115 120 125
64 Thr Gly Thr Val Ala Arg Leu Asp Gly Cys Glu Leu Phe Gly Tyr Pro
65 130 135 140
68 Val Lys Asp Ala His Arg Tyr Gly Val Gly Glu Ile Asp Ser Gly Gly
69 145 150 155 160
72 Arg Leu Leu Ser Leu Glu Glu Lys Pro Arg Arg Pro Arg Ser Asn Leu
73 165 170 175
76 Ala Val Thr Gly Leu Tyr Leu Tyr Thr Asn Asp Val Val Glu Ile Ala
77 180 185 190

```

## RAW SEQUENCE LISTING

DATE: 06/22/2006

PATENT APPLICATION: US/10/580,872

TIME: 12:53:02

Input Set : A:\SEQUENCE LISTING.txt

Output Set: N:\CRF4\06222006\J580872.raw

```

80 Arg Thr Ile Ser Pro Ser Ala Arg Gly Glu Leu Glu Ile Thr Asp Val
81      195      200      205
84 Asn Lys Val Tyr Leu Glu Gln Gly Arg Ala Arg Leu Thr Glu Leu Gly
85      210      215      220
88 Arg Gly Phe Ala Trp Leu Asp Met Gly Thr His Asp Ser Leu Leu Gln
89 225      230      235      240
92 Ala Gly Gln Tyr Val Gln Leu Leu Glu Gln Arg Gln Gly Glu Arg Ile
93      245      250      255
96 Ala Cys Ile Glu Glu Ile Ala Met Arg Met Gly Phe Ile Ser Ala Glu
97      260      265      270
100 Gln Cys Tyr Arg Leu Gly Gln Glu Leu Arg Ser Ser Ser Tyr Gly Ser
101      275      280      285
104 Tyr Ile Ile Asp Val Ala Met Arg Gly Ala Ala Ala Asp Ser Arg Ala
105      290      295      300
108 Gln
109 305
112 <210> SEQ ID NO: 2
113 <211> LENGTH: 303
114 <212> TYPE: PRT
115 <213> ORGANISM: Streptomyces fradiae
117 <400> SEQUENCE: 2
119 Met Asn Asp Arg Pro Arg Arg Ala Met Lys Gly Ile Ile Leu Ala Gly
120 1      5      10      15
123 Gly Ser Gly Thr Arg Leu Arg Pro Leu Thr Gly Thr Leu Ser Lys Gln
124      20      25      30
127 Leu Leu Pro Val Tyr Asp Lys Pro Met Ile Tyr Tyr Pro Leu Ser Val
128      35      40      45
131 Leu Met Leu Ala Gly Ile Arg Glu Ile Gln Ile Ile Ser Ser Lys Asp
132      50      55      60
135 His Leu Asp Leu Phe Arg Ser Leu Leu Gly Glu Gly Asp Arg Leu Gly
136 65      70      75      80
139 Leu Ser Ile Ser Tyr Ala Glu Gln Arg Glu Pro Arg Gly Ile Ala Glu
140      85      90      95
143 Ala Phe Leu Ile Gly Ala Arg His Ile Gly Gly Asp Asp Ala Ala Leu
144      100      105      110
147 Ile Leu Gly Asp Asn Val Phe His Gly Pro Gly Phe Ser Ser Val Leu
148      115      120      125
151 Thr Gly Thr Val Ala Arg Leu Asp Gly Cys Glu Leu Phe Gly Tyr Pro
152      130      135      140
155 Val Lys Asp Ala His Arg Tyr Gly Val Gly Glu Ile Asp Ser Gly Gly
156 145      150      155      160
159 Arg Leu Leu Ser Leu Glu Glu Lys Pro Arg Arg Pro Leu Glu Pro Gly
160      165      170      175
163 Arg His Arg Leu Tyr Leu Tyr Thr Asn Asp Val Val Glu Ile Ala Arg
164      180      185      190
167 Thr Ile Ser Pro Ser Ala Arg Gly Glu Leu Glu Ile Thr Asp Val Asn
168      195      200      205
171 Lys Val Tyr Leu Glu Gln Gly Arg Ala Ala His Gly Ala Gly Ala Val
172      210      215      220

```

## RAW SEQUENCE LISTING

DATE: 06/22/2006

PATENT APPLICATION: US/10/580,872

TIME: 12:53:02

Input Set : A:\SEQUENCE LISTING.txt

Output Set: N:\CRF4\06222006\J580872.raw

```

175 Val Ala Trp Leu Asp Met Gly Thr His Asp Ser Leu Leu Gln Ala Gly
176 225                230                235                240
179 Gln Tyr Val Gln Leu Leu Glu Gln Arg Gln Gly Glu Arg Ile Ala Cys
180                245                250                255
183 Ile Glu Glu Ile Ala Met Arg Met Gly Phe Ile Ser Ala Glu Gln Cys
184                260                265                270
187 Tyr Arg Leu Gly Gln Glu Leu Arg Ser Ser Ser Tyr Gly Ser Tyr Ile
188                275                280                285
191 Ile Asp Val Ala Met Arg Gly Ala Ala Ala Asp Ser Arg Ala Gln
192                290                295                300
195 <210> SEQ ID NO: 3
196 <211> LENGTH: 333
197 <212> TYPE: PRT
198 <213> ORGANISM: Streptomyces fradiae
200 <400> SEQUENCE: 3
202 Met Arg Val Leu Val Thr Gly Gly Ala Gly Phe Ile Gly Ser His Phe
203 1                5                10                15
206 Thr Gly Gln Leu Leu Thr Gly Ala Tyr Pro Asp Leu Gly Ala Thr Arg
207                20                25                30
210 Thr Val Val Leu Asp Lys Leu Thr Tyr Ala Gly Asn Pro Ala Asn Leu
211                35                40                45
214 Glu His Val Ala Gly His Pro Asp Leu Glu Phe Val Arg Gly Asp Ile
215                50                55                60
218 Ala Asp Gln Ala Leu Val Arg Arg Leu Met Glu Gly Val Gly Leu Val
219 65                70                75                80
222 Val His Phe Ala Ala Glu Ser His Val Asp Arg Ser Ile Glu Ser Ser
223                85                90                95
226 Glu Ala Phe Val Arg Thr Asn Val Glu Gly Thr Arg Val Leu Leu Gln
227                100               105               110
230 Ala Ala Val Asp Ala Gly Val Gly Arg Phe Val His Ile Ser Thr Asp
231                115               120               125
234 Glu Val Tyr Gly Ser Ile Ala Glu Gly Ser Trp Pro Glu Asp His Pro
235                130               135               140
238 Leu Ala Pro Asn Ser Pro Tyr Ala Ala Thr Lys Ala Ala Ser Asp Leu
239 145                150               155               160
242 Leu Ala Leu Ala Tyr His Arg Thr Tyr Gly Leu Asp Val Arg Val Thr
243                165               170               175
246 Arg Cys Ser Asn Asn Tyr Gly Pro Arg Gln Tyr Pro Glu Lys Ala Val
247                180               185               190
250 Pro Leu Phe Thr Thr Asn Leu Leu Asp Gly Leu Pro Val Pro Leu Tyr
251                195               200               205
254 Gly Asp Gly Gly Asn Thr Arg Glu Trp Leu His Val Asp Asp His Cys
255                210               215               220
258 Arg Gly Val Ala Leu Val Ala Ala Gly Gly Arg Pro Gly Val Ile Tyr
259 225                230               235               240
262 Asn Ile Gly Gly Gly Thr Glu Leu Thr Asn Ala Glu Leu Thr Asp Arg
263                245               250               255
266 Ile Leu Glu Leu Cys Gly Ala Asp Arg Ser Ala Val Arg Arg Val Ala
267                260               265               270

```

## RAW SEQUENCE LISTING

DATE: 06/22/2006

PATENT APPLICATION: US/10/580,872

TIME: 12:53:02

Input Set : A:\SEQUENCE LISTING.txt

Output Set: N:\CRF4\06222006\J580872.raw

```

270 Asp Arg Pro Gly His Asp Arg Arg Tyr Ser Val Asp Thr Thr Lys Ile
271      275      280      285
274 Arg Glu Glu Leu Gly Tyr Ala Pro Arg Thr Gly Ile Thr Glu Gly Leu
275      290      295      300
278 Ala Gly Thr Val Ala Trp Tyr Arg Asp Asn Arg Ala Trp Trp Glu Pro
279 305      310      315      320
282 Leu Lys Arg Ser Pro Gly Gly Arg Glu Leu Glu Arg Ala
283      325      330
286 <210> SEQ ID NO: 4
287 <211> LENGTH: 333
288 <212> TYPE: PRT
289 <213> ORGANISM: Streptomyces fradiae
291 <400> SEQUENCE: 4
293 Met Arg Val Leu Val Thr Gly Gly Ala Gly Phe Ile Gly Ser His Phe
294 1      5      10      15
297 Thr Gly Gln Leu Leu Thr Gly Ala Tyr Pro Asp Leu Gly Ala Thr Arg
298      20      25      30
301 Thr Val Val Leu Asp Lys Leu Thr Tyr Ala Gly Asn Pro Ala Asn Leu
302      35      40      45
305 Glu His Val Ala Gly His Pro Asp Leu Glu Phe Val Arg Gly Asp Ile
306      50      55      60
309 Ala Asp His Gly Trp Trp Arg Arg Leu Met Glu Gly Val Gly Leu Val
310 65      70      75      80
313 Val His Phe Ala Ala Glu Ser His Val Asp Arg Ser Ile Glu Ser Ser
314      85      90      95
317 Glu Ala Phe Val Arg Thr Asn Val Glu Gly Thr Arg Val Leu Leu Gln
318      100     105     110
321 Ala Ala Val Asp Ala Gly Val Gly Arg Phe Val His Ile Ser Thr Asp
322      115     120     125
325 Glu Val Tyr Gly Ser Ile Ala Glu Gly Ser Trp Pro Glu Asp His Pro
326      130     135     140
329 Val Ala Pro Asn Ser Pro Tyr Ala Ala Thr Lys Ala Ala Ser Asp Leu
330 145     150     155     160
333 Leu Ala Leu Ala Tyr His Arg Thr Tyr Gly Leu Asp Val Arg Val Thr
334      165     170     175
337 Arg Cys Ser Asn Asn Tyr Gly Pro Arg Gln Tyr Pro Glu Lys Ala Val
338      180     185     190
341 Pro Leu Phe Thr Thr Asn Leu Leu Asp Gly Leu Pro Val Pro Leu Tyr
342      195     200     205
345 Gly Asp Gly Gly Asn Thr Arg Glu Trp Leu His Val Asp Asp His Cys
346      210     215     220
349 Arg Gly Val Ala Leu Val Gly Ala Gly Gly Arg Pro Gly Val Ile Tyr
350 225     230     235     240
353 Asn Ile Gly Gly Gly Thr Glu Leu Thr Asn Ala Glu Leu Thr Asp Arg
354      245     250     255
357 Ile Leu Glu Leu Cys Gly Ala Asp Arg Ser Ala Leu Arg Arg Val Ala
358      260     265     270
361 Asp Arg Pro Gly His Asp Arg Arg Tyr Ser Val Asp Thr Thr Lys Ile
362      275     280     285

```

## RAW SEQUENCE LISTING

DATE: 06/22/2006

PATENT APPLICATION: US/10/580,872

TIME: 12:53:02

Input Set : A:\SEQUENCE LISTING.txt

Output Set: N:\CRF4\06222006\J580872.raw

```

365 Arg Glu Glu Leu Gly Tyr Ala Pro Arg Thr Gly Ile Thr Glu Gly Leu
366      290                      295                      300
369 Ala Gly Thr Val Ala Trp Tyr Arg Asp Asn Arg Ala Trp Trp Glu Pro
370 305                      310                      315                      320
373 Leu Lys Arg Ser Pro Gly Gly Arg Glu Leu Glu Arg Ala
374      325                      330
377 <210> SEQ ID NO: 5
378 <211> LENGTH: 2160
379 <212> TYPE: DNA
380 <213> ORGANISM: Streptomyces eurythermus
382 <400> SEQUENCE: 5
383 ggcatgcctt cgggggtgtgc ggcggcgccct cagagcgtgg ccagtacctc gtgcagggcc      60
385 gcgatcacct tgtcctgtac gtcgggcgcg agccccgggt acatcggcag cgagaagatc      120
387 tcgtccgcca gccgctccgt caccggcgagc gagcccttgg cgtaccccag gtgcgcgaag      180
389 cccgtcatgg tgtgcacggg ccacgggtaa ctgatgttga gcgagatccc gtacgacttg      240
391 agcgctcga tgatgtcgtc ccggcgcggg tggcggaaga cgtacacgta atacacgtgg      300
393 tcgttgccct cggtgacgga cggcagcacc agggcgccgg ggcccgtcag gttcgcgagt      360
395 ccttcggcgt aacgcggggc gaccgcgcgc cggccctcga tgtagcggtc gaggcgggtg      420
397 agcttgccgc gcaggatctc cgcctgcacc tcgtcgagcc ggctgttgtg gccgggcgtc      480
399 tgcacgacgt agtacacgtc ctccatgccg tagtagcgca gccggcgagc cgcacggtcg      540
401 acgtcccgct cgtcggtcag caccggcccc cgtcgccgt acgcaccgag gaccttcgtc      600
403 gggtagaacg agaaggcggg ggcgtcgccc agcgtgccgg ccagctcgcc gtgggtggcg      660
405 gcaccgtgcg cctggggcgca gtcctccagc accaccaggc cgtgctgctc ggccaggggc      720
407 cgcaagggcg ccatgtcgac gcaactgccg tacagggtga ccggcagcag ggcccttcgtg      780
409 cgcgggggtga tgacgtccgc gacctggctg gtgtccatga ggtggctcctc ggcgcggaag      840
411 tcgacgaaga cgggcgtggc accggtgccg tcgatggcca ccaccgtcgg cgcggccgtg      900
413 ttggagacgg tgacgacctc gtcccccggg cccacccga gcgcctgcag acccagcttg      960
415 acggcggttg tgccgttgtc gacaccgcgc cagtggcgca ggccgtggta gtccgcgaac      1020
417 tccttctcga acccgctcac gctggggccg aggaaccaact gcccggaagg gaagacggtc      1080
419 tcgacggcgt cgaggaggtc cgcgcgttcg ttctgggtatt ccgccaggta gtcccagacg      1140
421 taggtagtca cggagagctc aacctccaga gtgtttcgat ggggtgggtg gaagccgggtg      1200
423 cgcgcggacc aggtcgtgcc agcagtcgcg gaccgactcc cgcagcgaac ggcgcgggtg      1260
425 ccagcccagc agggcgcgcg ccgcgcgggt gtcgacccgc agccagtcct cccggtgcc      1320
427 gggagcccgg cccggagccg ggcgctccac caccgcgcgc ggaatgccgc tcgcctcgat      1380
429 gaacaggccg accaggtcgc ggacggcgac cgcctcgccc cgcgcgatgc cgacggcgac      1440
431 cgggacggcc ggtgcgcggg cggcgggcac gacggcgtcg gccacgtccc gcacatcgac      1500
433 gtagtcccgg tgcgcgcgca gccgggacag ttccacgacg gcctccgcac ccgtcccggc      1560
435 ggccgccagc agccgctcgg cgacctggcc cagcagactg atccgcgggg tgccggggcc      1620
437 cgacacgttg gacacccgta gcaccacacc gtcgacccac ccgcccaggg tgcccccgag      1680
439 caccgcctcg ctggcgggcg gcttgcctct gccgtacgcc gtgtccgggc gcggtacggc      1740
441 gtcggcgccc accgaaccgc cgggcgtcac cgggccgtac tccagtaccg agccgaggtg      1800
443 gaccagccgc ggccgcgcgg acatcagcgc cagcgccctc agcaggcgca gcgtgggcac      1860
445 cgcggtggcg gaccacatct gctcgtcggg acggccccag atgcttccga cggagttgac      1920
447 gatcgtgtcc ggacgtcccg cgtccagggg ggcggccagc gccgcgggat ccgtaccggc      1980
449 caggtccagg gtgacgcagc ggtacggcat cggctcctcg ggcggggcgg ggccccacc      2040
451 caccacgtca cggccccgcg cggcgaaacg cgcgcacaca tgccggccga cgtaccggc      2100
453 gccgcccagg accacgacgc tgccactgcc actgcccgcg ggcacccgat cgttcacccat      2160
456 <210> SEQ ID NO: 6
457 <211> LENGTH: 4461

```

RAW SEQUENCE LISTING ERROR SUMMARY      DATE: 06/22/2006  
PATENT APPLICATION:    US/10/580,872      TIME: 12:53:03

Input Set : A:\SEQUENCE LISTING.txt  
Output Set: N:\CRF4\06222006\J580872.raw

Invalid <213> Response:

Use of "Artificial" only as "<213> Organism" response is incomplete,  
per 1.823(b) of New Sequence Rules. Valid response is Artificial Sequence.

Seq#:16,17,18,19,20,21,22,23,24,25,26,27,28,29,30,31,32,33,34,35,36,37,38,39

Seq#:40,41,42,43,44,45,46,47,48,49,50,51,52,53,54,55,56,57

VERIFICATION SUMMARY

DATE: 06/22/2006

PATENT APPLICATION: US/10/580,872

TIME: 12:53:03

Input Set : A:\SEQUENCE LISTING.txt

Output Set: N:\CRF4\06222006\J580872.raw

L:12 M:270 C: Current Application Number differs, Replaced Current Application Number

L:13 M:271 C: Current Filing Date differs, Replaced Current Filing Date